



International Comparative Jurisprudence



COMPATIBILITY OF BUSINESS MODELS OF ESTONIAN LANGUAGE TECHNOLOGY ENTREPRENEURS WITH REGULATORY FRAMEWORK ON THE USE OF DIGITAL CONTENT

Irene Kull¹, Kadri Vider², Tõnis Mets³, Aleksei Kelli⁴

University of Tartu, Estonia

E-mail: irene.kull@ut.ee; kadri.vider@ut.ee; tonis.mets@ut.ee; aleksei.kelli@ut.ee

Krister Lindén⁵

University of Helsinki, Finland

E-mail: krister.linden@helsinki.fi

Received: 12 January 2020; accepted: 19 May 2020

DOI: <http://dx.doi.org/10.13165/j.icj.2020.06.006>

Abstract. Language technology provides several possibilities to commercialise collected language resources (data) in the form of providing access to databases, dictionaries, translation, text analysis and localisation services, storage of documents and personal language data, software, and other types of digital content. This article focuses on the contractual relationships in selling language technology products and providing services to the customers, and explores the entrepreneurial perspective on the supply of digital content to compare its compatibility with the existing regulatory framework. For comparison, the authors have chosen Estonian companies specialising in language technology. The main conclusions concerning the entrepreneurial perspective are based on a study of homepages and personal interviews. The compatibility between the business model and the regulatory framework is analysed based on the case study and Estonian legislation. The authors aim to outline how language technology entrepreneurs in Estonia themselves conceptualise the digital content they are commercialising, the nature of their business models, and entrepreneurial practice (business model). The authors take into account the applicable regulatory framework (protection of intellectual property, contract law, proprietary relations), and also analyse the contracting method, contractual remedies, and need for an objective neutrality in legal regulation.

Keywords: digital content, digital services, language technology, intellectual property rights, business models

Introduction

Digital content is a complex phenomenon. It can be conceptualised from different perspectives. In a technological context, it covers data in digital form (e.g., ebooks, software, applications, games, videos, music, etc.). From a legal point of view, digital content as an object of a legal transaction is protected by different rights such as intellectual property, data protection, property rights, and so forth. These rights, however, often remain invisible, since the main interest of the customer is usually to buy digital products or services (see Petrusson, 2011, p. 79). From an entrepreneurial perspective, digital content constitutes a business offering.

The authors have chosen Estonian companies specialising in language technology⁶ for further scrutiny. This choice

¹ PhD, professor of civil law at School of Law, University of Tartu.

² Research fellow in language technology at the Institute of Computer Science, University of Tartu.

³ PhD, professor at the School of Economics and Business Administration, University of Tartu.

⁴ PhD, professor of intellectual property law at the School of Law, University of Tartu.

⁵ PhD, research director of language technology at University of Helsinki.

⁶ Language technology (LT) can be defined as ‘an interdisciplinary field integrating information technology and linguistics which is concerned with developing language software and language resources necessary for the computational processing of

is conditioned by the fact that commercialised digital content and business models may vary depending on what digital content is supplied or what digital service is offered. The authors aim to combine traditional legal research methods with an empirical study considering the regulatory framework of digital content. For mapping the realm of language technology, the comparative case study (Yin, 2009) method was used alongside the collection of data. Collection of data means studying the structure of digital language resources (language data), involving multiple levels of analysis (product, market, industry; technology; institutions/actors; activities/events, and so forth), considering the structural embeddedness of language resource elements (variables); and employing an eclectic combination of data about the (tangible and intangible) elements of language resources, the commercial and legal regulatory framework, and their linkages.

The underlying feature of the current study is that language technologies encompass a wide range of products and services from electronic dictionaries to speech synthesizers. Consequently, an empirical study should cover a variety of the following aspects: language resources, learning and training tools, creation of digital content (of a product/service), commercialisation and relevant intellectual property rights (digital content is usually protected by copyright), and the legal rights and obligations of the parties.

This article outlines: (1) how entrepreneurs conceptualise the digital content they are commercialising and the nature of their business model; (2) the reliance on copyright protection in business models of language technology companies; (3) the legal meaning of digital content, the qualification of contractual relations, and how legal requirements concerning the information obligations are followed in practice; and (4) some conclusions concerning the conformity requirements and remedies. The authors analyse contractual relationships where ownership of the digital content is transferred, or where digital content is provided for use or as a service for counter-performance (payment) or for free. This brings us to the mapping of business models in the language technology business sector and gives some general summaries about the extent to which legal regulation (contract law and intellectual property law) is taken into account in the business models used by companies.

The authors draw on and develop further their previous research covering language resources and technology (Kelli et al., 2015; Kelli et al. 2019a; Kelli et al. 2019b).

1. Case Study on Business Models

The empirical study in this contribution involved mapping how language technology companies describe their product/service, the protection of their digital content, their target markets, their channels of distribution, and buyers' rights to the digital content of a product/service. Webpage and article searches were also conducted. For the collection of data from companies, a written interview form was used. Data analysis includes operations for systematisation, pattern recognition, and business model data interpretation of interviews (Ghauri, 2004, pp. 109–124) as presented in Table 1.

Table 1. *Case Study Analysis*

The technique for case study analysis	Explanation
Product/service	Description, features, and facts organised by theoretical concepts
Coding	Sorting data according to concepts and themes
Clustering	Categorising sorted data according to the patterns of digital content and models
Pattern recognition	Comparison of the models' and observed digital content patterns
Interpretation	Interpreting findings of aspects of digital content and business models

human language'. See the National Programme for Estonian Language Technology (NPELT) 2011–2017 (n.d.), confirmed by the Minister of Education and Research 25 January 2011, Directive No. 71. Retrieved 1 May 2020. <https://www.keeletehnoloogia.ee/en/npelt-text/view>

A framework for the mapping of variables is divided into four parts: (1) digital content of products and services; (2) ownership of digital content of products and services; (3) the questions related to change in digital content of products; and (4) identification of the main patterns of treating digital content and the commercialisation of language resources by language technology companies. The practical implementation of these patterns is discussed in the next sections. Aspects of the language ecosystem, including relations between research institutions, language technology companies, and customers/users of digital content, are the key drivers behind language technology development.⁷

The sample of the study was selected from language technology companies in Estonia. As there are only a small number of companies operating in the field, an expert method was used to ensure that the sampling covered the entire population of companies in the sector. The main criteria included their reliance on language technology development and their offering of these services on the market.

The sample of language technology companies in Estonia includes companies of different size and turnover. Information about the companies and entrepreneurs was collected from the transcriptions of interviews,⁸ companies' webpages, annual reports in the Commercial Registry,⁹ LinkedIn, and the Estonian Research Information System.¹⁰ The data based on these sources are not specifically referred to unless used in direct citations. The short company profiles below testify to the diverse nature of the language technology business in Estonia.

TEXTA was founded in 2016 as a subsidiary of Software Technology and Applications Competence Centre (hereinafter STACC).¹¹ The company is offering TEXTA full search, TEXTA LAW, and TEXTA Toolkit services. These services enable the conducting of real-time searches from datasets and the combining of the results across different data fields to identify multi-word semiotic units from the text corpus, to identify synonymous word forms, to form concepts, to create limited grammars of languages, and for the automatic categorisation of documents. Tilde Eesti was founded in 2000 as a subsidiary of Tilde SIA (Latvia).¹² Tilde was awarded a subsidy for the development of machine translation from the ME of language technology of Estonia. The company provides machine translation, virtual assistants, translation and localization services, technology services, and language technology development are only secondary to their core activities. The company exports its services to Latvia and other countries in Europe, and as such the share of the Estonian accounts amounts to 13.8%. Filosoft was founded in 1993 by academicians employed mainly by the university. It is a small business venture of only three persons operating on a part-time basis.¹³ Filosoft offers free SaaS (Software as a Service) language help and software, including an Estonian speller, grammar, hyphenator and thesaurus, etc. Premium products are the speller, hyphenator, and the thesaurus for Estonian, the English–Estonian dictionary, and the speller and hyphenator for Latvian. Additional income is gained from public target funding projects.

The data regarding the level of knowledge of the legal aspects of their products and services was collected from the companies by telephone interviews. The answers are systematically displayed in Annex 1 to provide an

⁷ The research and development program Estonian Language Technology 2018–2027 provides that '[the Estonian] state supports a language technology in the fields where it is not always profitable for the private sector to take on the risks associated with the development of technology for a language with a small number of speakers – as a small number of speakers also means a small market'.

⁸ First telephone interviews were conducted by Kadri Vider: Filosoft on 17 May 2017, Tilde on 19 May 2017, and TEXTA on 22 May 2017. The re-survey was conducted in 2019.

⁹ Retrieved 1 May 2020. <https://ariregister.rik.ee/index?lang=eng>

¹⁰ Retrieved 1 May 2020. <https://www.etis.ee/Portal/Persons/Index?#>

¹¹ Retrieved 1 May 2020. <https://www.texta.ee/home>

¹² Retrieved 1 May 2020. <https://www.tilde.com>

¹³ Retrieved 1 May 2020. https://www.filosoft.ee/index_en.html

overview of the business models and legal approaches of the abovementioned Estonian language technology companies.

2. Copyright Protection of Language Technology

Language technology has two main branches: language resources and software.¹⁴ Language resources are copyright protected databases containing language data. Language data itself is often covered with copyright, related rights, and personal data rights. This article focuses on language software (computer programs) since the authors have already previously addressed the issues relating to language resources (Kelli et al., 2016, pp. 81–98; Kelli, et al., 2015, pp. 13–24; Kelli et al., 2012, pp. 40–48).

Language technology is the technology for processing language. In most cases, such technology is embodied in a piece of software and more seldom in an explicit device. It used to be the case that all language technology software was developed like any other computer program: as a tailor-made piece of software to solve a particular problem, e.g. a speller for Estonian knowing about all the words and their inflections in Estonian encoded in the software. As science progressed, it became evident in the 1980s that words and their inflections could be encoded along with similar principles in different languages, and that spelling correction consisted of similar language-independent actions. This gave rise to the idea that language-specific information could be encoded in a database of language-specific facts, and that the language processing software could be made generic and language-independent. By changing the language-specific database, the software could be made to process a different language. This opened the previously local language-specific markets of language technology for international competition, which largely took over many of the local markets in the beginning of the 2000s starting with the more profitable languages with many speakers, because collecting databases of language-specific facts was a hand-made effort (Kelli et al., 2012, pp. 40–48). By the end of the 2010s, machine learning had entered into the picture and taken over much of the effort to collect language-specific facts into databases. When collecting such language-specific databases, the human effort was mainly devoted to guiding and initiating the mechanical extraction process and fine-tuning the end-result. This also meant that language technology markets with fewer speakers such as the Estonian market began facing international competition.

With the advent of more stable computer networks, the demand for software as a service is gaining ground. Instead of installing a piece of software on a computer, it is possible to provide access to a computer server providing the same functionality on demand. With time, this will shift the focus of the language technology entrepreneurs from selling a copy of a piece of software to selling access to an electronic service.

Software protection relies on copyright. According to the Directive 2009/24/EC on the legal protection of computer programs (hereinafter Software Directive), computer programs are protected by copyright as literary works (Art. 1, Directive 2009/24/EC). The Estonian Copyright Act (hereinafter CA), which is harmonised with the Software Directive, provides that copyright protection applies to the expression in any form of software (§ 4 [3] 3), CA).

The Court of Justice of the European Union (CJEU) has also provided some guidelines how to interpret the Software Directive. It is necessary to define the scope of software protection since there are special regulations (*lex specialis*) covering software at the EU level (e.g. digital exhaustion). For instance, the CJEU has specified that a graphic user interface (GUI) is not a form of expression of a computer program within the meaning of the

¹⁴ According to NPELT, 'Language software involves methods for processing language materials, algorithms, and computer programmes; and is the basis for language technology application systems. It is useful to distinguish between speech technology (e.g. speech recognition and speech synthesis) and the technology for processing written text (e.g. morphological, syntactic, and semantic analysis), i.e. language technology in its more restricted sense. Language resources are electronic databases that are used to develop language software: corpora (collections of speech signals and texts), electronic dictionaries and databases'.

Software Directive. GUIs, however, can be protected by copyright in case the GUI is its author's intellectual creation (Case C-393/09). In other words, the CJEU took the view that a GUI could be copyright-protected, but it does not fall within the regulation of the Software Directive.

The CJEU has also explained that:

neither the functionality of a computer program nor the programming language and the format of data files used in a computer program in order to exploit certain of its functions constitute a form of expression of that program and, as such, are not protected by copyright in computer programs for the purposes of that directive (Case C-406/10).

This approach is compatible with the requirement outlined in the Software Directive: that protection applies to expression, and that ideas and principles, which underlie software, are not protected (Art. 1[2], Directive 2009/24/EC).

Language technology companies sell, provide access to digital content, or offer digital services based on language technology. This raises the issue of what the customer can do with the available content under the current copyright framework at the EU level. For example, in its seminal *UsedSoft* decision, the CJEU has acknowledged transferrable ownership of a digital copy in its assertion that:

the right of distribution of a copy of a computer program is exhausted¹⁵ if the copyright holder who has authorised, even free of charge, the downloading of that copy from the internet onto a data carrier has also conferred, in return for payment of a fee intended to enable him to obtain a remuneration corresponding to the economic value of the copy of the work of which he is the proprietor, a right to use that copy for an unlimited period (Case C-128/11).

The CJEU has later reaffirmed that 'the initial acquirer of a copy of a computer program accompanied by an unlimited user licence is entitled to resell that copy and his licence to a new acquirer' (Case C-166/15).

These cases, however, concern software, and they are not directly applicable to other digital content (e.g. ebooks, music, etc.). The CJEU held in its *Allposters* case C-419/13 that:

the rule of exhaustion of the distribution right ... does not apply in a situation where a reproduction of a protected work, after having been marketed in the European Union with the copyright holder's consent, has undergone an alteration of its medium, such as the transfer of that reproduction from a paper poster onto a canvas, and is placed on the market again in its new form.

At the same time, the *VOB* case C-174/15 confirms that the concept of 'lending' covers the lending of a digital copy of a book. In the *Tom Kabinet* case, the court decided that:

The supply to the public by downloading, for permanent use, of an ebook is covered by the concept of 'communication to the public' and, more specifically, by that of 'making available to the public of [authors]' works in such a way that members of the public may access them from a place and at a time individually chosen by them (C-263/18).

Legal commentators have pointed out that:

¹⁵ According to the Software Directive Art. 4 (2): 'The first sale in the Community of a copy of a program by the right-holder or with his consent shall exhaust the distribution right within the Community of that copy, with the exception of the right to control further rental of the program or a copy thereof'.

the scope of the exhaustion principle differs depending upon whether the work falls to be protected by the Software Directive or the InfoSoc Directive... This is an unsatisfactory position for many reasons, not least because certain types of complex works could potentially be protected by both the InfoSoc Directive and the Software Directive (Ohta, 2015).

To sum up, there is a lack of clarity regarding the ownership and transferability of a digital copy (for further discussion see Oprysk 2020). Therefore, it is crucial to learn how entrepreneurs themselves define their business models.

The digital content supplied by the surveyed companies is copyright protected. TEXTA has three types of business offers: 1) TEXTA Law, a searchable database, is protected by copyright as a database; 2) TEXTA Toolkit is software protected by copyright; and 3) Text analysis as a service which is also protectable by copyright. To sum up, TEXTA relies on copyright. Their content is licensed, and a dual licensing scheme is applied so that research use is encouraged. Except for TEXTA Law, TEXTA consents to customers becoming the owners of digital content. This means that TEXTA accepts digital exhaustion of the content delivered and customers can change and resell it. Tilde Eesti delivers localisation and translation services to customer. The results are copyright protected. The customer becomes the owner of the translation and is allowed to change and resell it. FiloSoft sells software for text analysis purposes (i.e. spellers). It relies on copyright protection and a technological progress. Technology becomes obsolete, so the product can in practice be used only for a limited period. The source code is not public, and the software is non-exclusively licensed. FiloSoft has no objections to a customer transferring a copy of the software bought. However, a separate contract is needed to modify the software.

Based on the case studies, we can conclude that the surveyed companies operate within the current copyright framework. At the same time, copyright is just one element among several others in their business environment. There is no more intensive attention given to copyright issues in comparison to business considerations, customer relations, and technological progress, which are equally important in their activities.

The reliance on the copyright framework by the sample companies could be partly explained by the public funding source of their product and service development, as three of them have been supported by the National Programme for Estonian Language Technology 2011–2017, or the key personnel of the companies having prior connections with public institutions. A substantial share of public funding in their revenue and the infinitesimally small market of Estonian language resources both affect the commercial interests of these companies. Such factors have undoubtedly shaped their business models and related contractual patterns, as discussed below.

3. Legal Concept of Digital Content and Specification of Contractual Relations

Digital content is defined in the § 14¹(1)(3) of the Estonian Law of Obligations Act (hereinafter LOA) as data which are produced and supplied in digital form in accordance with the Art. 2 (11) of the Directive 2011/83 on consumer rights. Directive 2019/770 on certain aspects concerning contracts for the supply of digital content and digital services follows the definition of digital content used in the Directive 2011/83/EU of the European Parliament and of the Council of 25 October 2011 on consumer rights, and in addition includes the definition of digital services (Art. 2[1]). According to Directive 2019/770, a digital service is a service that allows the consumer to create, process, store, or access data in digital form, or a service that allows the sharing of or any other interaction with data in digital form uploaded or created by the consumer or other users of that service (Art. 2 [2], Directive 2019/770).

Digital content provided by the language technology companies is supplied in an intangible form, for example being downloaded, streamed, or accessed on the internet. In the legal sense it may be a contract on sale of goods, lease, licence agreement, contract for work, services, or some other type of contract (Spindler, 2016, p. 184 ff), as Estonian law does not automatically associate the concept of digital content with any particular type of contract. In any case, legal rules applicable to the specific contract shall be determined on the basis of the form of supply

of the digital content and services (rec 56, 57, Directive 2019/770).

The contract for the supply of digital content could be regarded as a sales contract under Estonian law if a seller undertakes to deliver an existing thing, a thing to be manufactured or produced, or a thing to be acquired in the future by the seller to the purchaser and to allow the transfer of ownership to the purchaser, and the purchaser undertakes to pay the purchase price for the thing to the seller in cash and to take delivery of the thing (§ 208 [1], LOA). An object of the sales contract can also be a right (§ 208 [3], LOA). Transfer of ownership is regulated under the Estonian Law of Property Act (1993).

If the customer obtains only the right to use the digital content, the rules on contract of licence (§ 368 ff, LOA) or contract for commercial rent (§ 339 ff, LOA) may apply (Sein, 2017, p. 106). Under Estonian law, by a licence agreement, one person (the licensor) grants another person (the licensee) the right to exercise the rights arising from intellectual property to the agreed extent and on the agreed territory, and the licensee undertakes to pay a fee (the licence fee) therefore (§ 368, LOA). If the licensee has been granted a non-exclusive licence agreement, the licensor may also exercise the right which is the object of the agreement, or grant the right of use to other persons besides the licensee. In the case of an exclusive licence agreement, the licensee is granted the rights arising from intellectual property to the agreed extent, and this precludes the right of use of other persons and the licensor to the same extent (§ 270, LOA). The contracts for the supply of digital content are rather deemed to be non-exclusive licence agreements. However, departing from the ‘classical approach’ of licensing of intellectual property rights applicable in most Member States, the contractual rights and obligations of the parties shall not be influenced (Loos et al., 2011; Guibault et al., 2010, p. 81 ff; rec 12, Directive 2019/770).

There is a possibility that parties conclude a commercial lease contract, under which one person (the commercial lessor) undertakes to grant the use of the object of the contract to another person (the commercial lessee) and enables the fruits received from the object of the contract to be enjoyed under the rules of regular management (§ 339 of LOA). It has to be added that the main purpose of the commercial lease is to gain profit (Paal, 2007, p. 145).

If digital content as a product or service will be developed according to the consumer's specifications as a tailor-made product, the provisions concerning the contract for work (§ 635, LOA) could be applicable. By a contract for work, one person undertakes to manufacture or modify a thing, or to achieve any other agreed result, by providing a service for payment (§ 635, LOA).

Finally, there is a possibility that the specific type of contract will not be defined, and contracts for the supply of digital content are regarded as *sui generis* contract (Spindler, 2016, pp. 185–186). Under LOA § 1(1) provisions of the general part of the LOA apply to all contracts which are not regulated by law but are not in conflict with the content and spirit of the law.

Whether the contract for the supply of digital content shall be regarded as a digital service, sales, or commercial lease agreement or *sui generis* contract, the customer's interest in obtaining digital content or services in accordance with the contract or law has to be protected (Ortiz & Viscasillas, 2012, p. 251; Jacquemin, 2017, pp. 27–38). Also, in case of a supply of digital content provided for free, e.g. in exchange for access to personal data, all remedies provided for in the law have to be available to the customer (rec 24, Art. 3[1], Directive 2019/770; Manko & Monteleone, 2017). There is a possibility that downloading will be treated as a service and a download itself as a ‘quasi-good’ meeting the requirements of the copyright (Reed, 2010, p. 248 ff; Case C-128/11, para. 47).¹⁶ The Directive 2019/770 approach to this issue seems to be negative while it does not differentiate between the rights concerning digital content supplied under licence agreements from those ‘sold’ in a traditional sense (rec 53, Directive 2019/770).

¹⁶ CJEU made it clear that by applying the principle of exhaustion to downloading and storing of software on customers' computers, there is no difference in which form the software was supplied.

Estonian language technology companies studied provide a wide range of digital services which can be qualified into several contractual types.¹⁷ For example, Tilde provides machine translation, terminology services, professional localisation, solutions for Baltic languages, virtual assistance, multilingual fonts and data library, e-services for developers.¹⁸ FiloSoft provides services for businesses by selling software, programs and providing access to freeware. TEXTA offers a platform as an environment for basic searches, prefix searches, and phrase searches from all indexed documents. The TEXTA LAW programme provides a tool for legal experts who need fast and modern access to public legal information, and the Texta toolkit includes the tools necessary for text analysis, e.g., text searcher, multiword expression miner, base lexicon miner and conceptualiser, grammar miner and classifier. The Texta toolkit also uses tailor-made solutions for the customer accessible as a cloud service or installed in the customer's infrastructure, e.g. analytical services. STACC provides text mining, machine learning, bioinformatics and other language technology services.

The services, which include creation, processing, or storage of digital data provided by the customer, and where data flows from the customer to the service provider, can be qualified as contracts for services or contracts for work. The use of digital content by the customer usually includes the creation of a text document using an online office package or the creation of documents and texts online and can be qualified as contracts for services. Rules on contracts for services also apply to the processing of digital data, even if some form of data will be created for future processing by the service provider (Maňko, 2016). Digital services also cover storage of digital data provided by the customer, e.g. cloud computing and services allowing sharing and any other interaction with data in digital form provided by other users of the service (cloud computing). If tailor-made digital content is offered, the contract can be treated as a contract for work. For example, subject to individual contracts, TEXTA offers individual creation of solutions for particular tasks. However, businesses do not accord too much importance to the types of contracts concluded. There are service providers who view themselves as sellers and, conversely, sellers who often define themselves as service providers.

FiloSoft offers programs used on the internet for free (Estonian HTML speller, morphanalysis of Estonian HTML, lemmatizer of the Estonian language, Estonian language synthesizer, bisector of the Estonian language, thesaurus of the Estonian language, HTML Hyperlinker, English–Estonian dictionary), software for payment, and free applications. The purchase of the software is contracted through the website. This requires entering the personal data of the buyer and the registration number of the software. By clicking on 'buy', the conditions of the contract appear on the screen.¹⁹ Tilde offers translation, localization, and terminology services for payment via individually negotiated contracts. Yet no information about the terms is available on the homepage except for the description of the service, technical requirements, and mode of payment.

Licensing is infrequently discussed on the websites of the businesses at issue. Digital content is commonly supplied to the customer under a licence entitling the customer to use the digital content under specific (often limited) circumstances or to obtain ownership. For example, Tilde offers service contracts for translation and editing of texts, under which the customer will maintain the ownership of the generated text; however, the software is not part of the deliverable to the customers. Customers of TEXTA become owners of the digital content except when using services of TEXTA Law, as the ownership of the database of court decisions is not a deliverable. FiloSoft relies on the term 'sale of licence' and provides the right to resell the purchased program; however, copies can only be made for personal use, but not for resale.

¹⁷ In the questionnaire there were no questions concerning the type of contracts concluded with the customer.

¹⁸ Description of the services is available at <https://www.tilde.com/products-and-services> Retrieved 1 May 2020.

¹⁹ The conditions of purchase are available only in Estonian. Retrieved 1 May 2020. <http://www.filosoft.ee/products/buy.php?productid=o2016>

Digital services are provided by the companies outlined in exchange for prior user registration, which can be carried out instantly by displaying user data and noting that the customer has been provided with the corresponding service conditions. In general, conditions can be distinguished by their user (terms and conditions) and privacy settings (privacy policy).

The degree of information available on the conclusion of the contract, contract terms, or the technical means for conclusion differ significantly. For example, Filosoft provides information about the programs needed to use the software and provides assurances that the software is functioning in conformity with the documentation supplied on the acquisition, and contains no intentionally installed obstacles which may impede or halt (cut off) the functioning of the software in the future. Filosoft also warns that the hardware or software updates made by a third party may affect the functioning of the software. Also, information about processing the order and payment is available. Filosoft tends to prefer one-time payments. The contract term can be copied from the homepage. Immediately after the payment is made in electronic form, the trader confirms the receipt of the order. If the customer prefers any other payment method (linked to e-bank), they can be contacted personally for the conclusion of the contract. In conclusion, Filosoft is keeping within the requirements of the law regarding contracting through a computer network.

The homepage of TEXTA provides information about the services, while the formation of the contract is personalised which means that information about the conditions of the contract is not publicly available. TEXTA provides services to businesses and consumers alike. TEXTA law offers a periodic payment system, whilst also using a one-time payment method for some services. The price depends on the mass of data owned, as opposed to the right of ownership obtained. Tilde uses one-time payments. However, if the service is to be provided under a long-term contract, the payments are periodic.

In conclusion, contracts used by language technology companies are different in their legal nature and level of individualization. Directive 2019/771 on certain aspects concerning contracts for the sale of goods seems to be a reasonable attempt to avoid fragmentation based on the regulation of information obligations being dependent on the way a contract is concluded (Lehmann, 2016, p. 763). This means that a customer enjoys the same rights regardless of how the contract is concluded (offline or online). There is still uncertainty in cases where the contract does not meet the characteristics of a sales contract but is a contract of services or another type of contract. Questions of ownership or licensing rights have been dealt with in the process of contracting only in rare cases. Several services are offered for free. If the service is provided for free, the amount of personal data is limited (name, address, phone or an email). However, there is no information about the use of data for economic purposes (Narciso, 2017). Given the requirements of the law, we may conclude that the businesses discussed, in general, provide the information required under the law depending on whether the contract is concluded online, individually, or services are provided for free.

4. Conformity Requirements and Remedies

Digital goods and services have to comply with the requirements provided for in the contract and law (Spindler, 2016, p. 198). Directive 2019/770 combines subjective and objective approaches providing that, to the extent that the contract does not stipulate where relevant, in a clear and comprehensive manner, the requirements for the digital content, the digital content shall be fit for the purposes for which the digital content of the same description would normally be used including its functionality, interoperability, and other performance features such as accessibility, continuity, and security (Art. 7 and 8, Directive 2019/770). The same approach can also be found in Estonian law. Paragraph 217 (1) of LOA provides that the goods delivered to a purchaser shall conform to the contract, in particular in respect of the quantity, quality, type, and description. Also, documents relating to a thing shall conform to the contract. In accordance to § 217(1)(2) of LOA, goods do not conform to a contract if the goods do not have the agreed characteristics, or are not fit for the particular purpose for which the purchaser needs them and of which the seller was or ought to have been aware at the time of entry into the contract if the purchaser could reasonably expect to rely on the professional skills or expertise of the seller, and in other cases for the

purposes for which such things would ordinarily be used. Also, the goods have to be free of rights of third parties which they may exercise with respect to the goods (§ 217 [2][3], LOA). The lack of conformity of goods arising from incorrect installation is deemed to be equal to the lack of conformity arising from the goods if the installation was carried out by the seller or at the responsibility of the seller (see § 217 [5], LOA).

Under Estonian law, the standard of the seller's liability is very high (Kalamees, 2011, p. 65; Kõve, 2008, p. 204; Saare et al., 2008, pp. 43–53). According to § 218 (1) of LOA, the seller is liable for any lack of conformity of goods which exists at the time when the risk of accidental loss of, or damage to, the goods passes to the purchaser even if the lack of conformity becomes apparent after that time (Värv & Karu, 2009, pp. 85–93; Zoll, 2016, p. 252; Smits, 2016, p. 10). The same standard could also be applied in the case of digital goods (Art. 10(1), Directive 2019/771). In consumer contracts, the agreement on the passing of risk does not influence the standard of liability provided for in the law (§ 218 [2], LOA). However, the concept of conformity of specific goods and services like digital content and digital services has to be interpreted subject to the purpose and content of every individual contract. Directive 2019/770 provides a list of the specific elements of digital content and services which have to be in conformity with the contract. The main problem seems to be how businesses will apply subjective and objective standards to the digital content and services they provide. It has to be mentioned that, under Estonian law, similar rules apply to the contracts for work (§ 641, LOA) concerning the conformity of final result of the promised work and liability of the contractor for non-conformity.

Language companies provide information about the compatibility of the software, but support and consultation regarding the use of different programs and tools, as well as regular system improvements with additional data and components, are offered. Exclusion of liability for non-conformity of digital content and for any direct or indirect patrimonial or non-patrimonial damage caused by the software can be found in some cases (Filosoft). Under Estonian law, the standard term which precludes the liability arising from the law of the party supplying the standard term or restricts such liability in the case where damage is caused intentionally or due to gross negligence is presumed to be unfair (§ 42 [3]1), LOA) and void in B2C contracts. The parties can agree in advance to preclude or restrict liability in the case of non-performance of an obligation (§ 106 [1], LOA). Agreements under which liability is precluded or restricted in the case of intentional non-performance or which allow the obligor to perform an obligation in a manner materially different from that which could be reasonably expected by the other party, or which unreasonably exclude or restrict liability in some other manner, are void (§ 106 [2], LOA).

Conclusions

Language technology companies in Estonia encompasses a wide range of products and services from electronic dictionaries to speech synthesisers. This study shows that language technology services are widely commercialized and several business models are used. Conceptualisation of the digital content from the legal perspective is not common, including copyright issues and questions relating to ownership of digital content if choosing a business model. However, as a rule, ownership of the digital content and subsequent right to resale are accepted. Numerous non-legal aspects, such as technology (e.g. technological obsolescence, and so forth) and the limited size of the market, are the main influencing forces in choosing business models. Even if digital content and digital services are provided from a distance, the conclusion of on-premises contracts might be required.

Contracts concluded by language technology companies can be classified under Estonian law as sales contracts and contracts for services. Some companies also provide digital content and services under the contract of licensing. The supply of digital content and services for free is not conceptualized as a contractual relationship. Finally, contract law and intellectual property law rules are playing limited roles for language technology companies when shaping their business models. Practical and legal problems with business models are likely to be expected in the future if language technology businesses' scant awareness of their rights and legal obligations towards their customers continues.

Annex 1. *Overview of the Business Models and Legal Approaches of the Sample of Estonian Language Technology Companies*

Company, founding year	TEXTA (spin-off of STACC), 2016	Tilde SIA, 1991/Tilde Eesti, 2000	Filosoft, 1993
Brief description of products and services	TEXTA full search, TEXTA LAW and TEXTA Toolkit services; automatic categorisation of documents	Machine translation	Premium products: (1) speller, hyphenator and thesaurus for Estonian, (2) English–Estonian dictionary, and (3) speller and hyphenator for Latvian
Target markets	Currently Estonia. Plans to move to Scandinavia, Germany, etc. Medium and large companies.	Global market. All customers who need translation or localisation services.	Global. IT companies, end users, other companies incorporating the product into their end products.
Distribution channel	Currently Facebook and direct contacts.	Business customer sales, conferences, fairs, internet sale, etc.	Direct sales and e-shop
Digital content and description of product/service	1) TEXTA Law is a web service enabling access to pre-processed textual data and an environment for text and data mining. Data is collected from public sources such as Estonian, EU, and German case law, decisions of the Consumer Disputes Committee, and Public Procurement Review Committee. 2) TEXTA Toolkit (text analysis tools as software). The customer inserts data or TEXTA helps to do it. 3) Text analysis as a service. If the customer requires a solution (e.g. to classify web posts) then TEXTA creates a model and develops an application programming interface (API) which the customer integrates into their system. The customer receives an API code and documentation.	1) Language technology: machine translation to individuals through browsers and to machines through API. 2) Localisation and translation services to business customers. Mainly translations into Estonian. Digital content covers the translation of the text delivered in a certain format (e.g. pdf) or structured machine-readable data (e.g. XML) compatible for the integration of the customer's system (e.g. Word document). Digital content does not cover software (e.g. used for the transformation of data formats, machine translation, etc.) which is meant for in-house use.	Software for text analysis, i.e. a speller. (e.g. speller for Estonian)
Protection of digital content	Software is protected by copyright. The dual licensing scheme is applied: research use is free (not limited to research institutions), commercial use requires payment.	Customer contracts relating to localisation and translation services determine the required level of data security of digital content. Data security measures encompass technical (securing servers), organisational (business	Software is protected by copyright. The protection is mainly based on technological process. Technology becomes obsolete and the product can be used for a limited period. The source code is not public. Decompilation is not possible.

	Customer contracts and security measures are stored on a server. Licence for a certain number of users. Based on API the customer decides who has access.	processes), and legal (instructions) levels.	
Payment for digital content and factors influencing price.	Payment for TEXTA Law is periodic, for other services a one-time payment. The amount of data owned and included in the environment by TEXTA could influence the price. More data, higher price. The issue of ownership of digital content does not have an impact on the price.	It is usually a one-time payment.	It is usually a one-time payment. Technical support is continuous if the software has deficiencies. The customer who bought the previous version of the software is entitled to buy the new version (e.g. compatible with new MS Office) for a reduced price. It is not possible to obtain software for development with a one-time payment.
Ownership of digital content	TEXTA has three types of services. In the case of TEXTA Law, the user does not become the owner of digital content. For other services, the customer become owners.	The customer becomes the owner of a translation.	The software is non-exclusively licensed. The customer can use the software. Organisations can buy multiple licences for reduced prices.
Right to transfer digital content	It is not allowed for TEXTA Law database. For other services the customer can resell the content.	The customer can transfer the translation.	The customer can transfer the copy bought. The customer is not allowed to make copies of the software for sale. Copying for personal use is permitted.
Alterations, modifications and adaptations to digital content	The customer has the right to change content (except TEXTA Law)	The customer has the right to change the content. There are no restrictions.	Contracts with IT developers allow occasional modification of the software to ensure compatibility with the existing systems. Since the software is user-friendly then the right to modify is limited.
Compatibility / incompatibility	Compatibility is guaranteed.	Compatibility is not an issue since the customer determines the format of the translation delivered (digital content).	Technical support is provided to assure compatibility. Obsolete software will be replaced with new one.

References

- Consumer Rights Act (2015). <http://www.legislation.gov.uk/ukpga/2015/15/contents/enacted>
- Copyright Act (1992). Entered into force on 12 December 1992; RT I 1992, 49, 615. <https://www.riigiteataja.ee/en/eli/524012017001/consolide>
- Directive 97/7/EC of the European Parliament and of the Council Text with EEA relevance. *Official Journal of the European Union*, L 304, 22.11.2011, pp. 64–88.
- Directive 2011/83/EU of the European Parliament and of the Council of 25 October 2011 on consumer rights, amending Council Directive 93/13/EEC and Directive 1999/44/EC of the European Parliament and of the Council and repealing Council Directive 85/577/EEC and Directive 2009/24/EC of the European Parliament and of the Council of 23 April 2009 on the legal protection of computer programs (Codified version) (Text with EEA relevance). *Official Journal of the European Union*, L 111, 5.5.2009, pp. 16–22.

- Directive 2019/770 of the European Parliament and the Council on 20 May 2019, regarding certain aspects concerning contracts for the supply of digital content and digital services (Text with EEA relevance). *Official Journal of the European Union*, L 136, 22.5.2019, pp. 1–27.
- Directive 2019/771 of the European Parliament and of the Council of 20 May 2019 on certain aspects concerning contracts for the sale of goods, amending Regulation (EU) 2017/2394 and Directive 2009/22/EC, and repealing Directive 1999/44/EC (Text with EEA relevance.) *Official Journal of the European Union*, L 136, 22.5.2019, pp. 28–50.
- Directive 2001/29/EC of the European Parliament and of the Council of 22 June 2001 on the harmonisation of certain aspects of copyright and related rights in the information society. *Official Journal of the European Union*, L 167, 22.6.2001, pp. 10–19.
- Ghauri, P. (2004). Designing and conducting case studies in international business research. In R. Marschan-Piekkari, & C. Welch (Eds.), *Handbook of qualitative research methods for international business*. Edward Elgar Publishing.
- Guibault, L., Helberger, N., Loos, M., Mak, C., Pessers, L., & van der Sloot, B. (2012). *Digital consumers and the law. Towards a cohesive European framework*. Wolters Kluwer.
- Jacquemin, H. (2017). Digital content and sales or service contracts under EU law and Belgian/French law. *Journal of Intellectual Property, Information Technology and Electronic Commerce Law*, 8(1), 27–38.
- Judgment of the Court (Fourth Chamber) of 22 January 2015 in Case C-419/13 Art & Allposters International BV v. Stichting Pictoright. ECLI:EU:C:2015:27.
- Judgment of the Court (Grand Chamber), 2 May 2012 in Case C-406/10 SAS Institute Inc. v. World Programming Ltd. ECLI:EU:C:2012:259.
- Judgment of the Court (Grand Chamber), 3 July 2012 in Case C-128/11 UsedSoft GmbH v. Oracle International Corp. ECLI:EU:C:2012:407.
- Judgment of the Court (Grand Chamber), 19 December 2019 in Case C-263/18, Nederlands Uitgeversverbond and Groep Algemene Uitgevers v. Tom Kabinet Internet BV and Others. Request for a preliminary ruling from the Rechtbank Den Haag. ECLI:EU:C:2019:1111.
- Judgment of the Court (Third Chamber), 10 November 2016 in Case C-174/15 Vereniging Openbare Bibliotheken v. Stichting Leenrecht. ECLI:EU:C:2016:856.
- Judgment of the Court (Third Chamber), 12 October 2016 in Case C-166/15 Aleksandrs Ranks and Jurijs Vasiļevičs v. Finanšu un ekonomisko noziegumu izmeklēšanas prokuratūra and Microsoft Corp. ECLI:EU:C:2016:762.
- Judgment of the Court (Third Chamber) of 22 December 2010 in Case C-393/09 Bezpečnostní softwarová asociace – Svaz softwarové ochrany v. Ministerstvo kultury. ECLI:EU:C:2010:816.
- Kalamees, P. (2011). Hierarchy of buyer's remedies in case of lack of conformity of the goods. *Juridica International*, 18, 63–72.
- Kelli, A., Lindén, K., Vider, K., Kamocki, P., Birštonas, R., Calamai, S., Labropoulou, P., Gavrilidou, M., & Straňák, P. (2019a). Processing personal data without the consent of the data subject for the development and use of language resources. In I. Skadina & M. Eskevich (Eds.), *Selected papers from the CLARIN Annual Conference 2018* (pp. 72–82). CLARIN Annual Conference 2018, Pisa, Italy, 8–10 October 2018. Linköping University Electronic Press. <https://www.ep.liu.se/ecp/article.asp?issue=159&article=008&volume>
- Kelli, A., Tavast, A., Lindén, K., Vider, K., Birštonas, R., Labropoulou, P., Kull, I., Tavits, G., & Värvi, A. (2019b). The extent of legal control over language data: The case of language technologies. In K. Simov & M. Eskevich (Eds.), *Proceedings of CLARIN Annual Conference 2019: CLARIN Annual Conference, Leipzig, Germany, 30 September – 2 October 2019* (pp. 69–74). CLARIN. <https://office.clarin.eu/v/CE-2019-1512-CLARIN2019-ConferenceProceedings.pdf>
- Kelli, A., Tavast, A., & Pisuke, H. (2012). Copyright and constitutional aspects of digital language resources: The Estonian approach. *Juridica International*, 19, 40–48.
- Kelli, A., Vider, K., Lindén, K. (2015). The regulatory and contractual framework as an integral part of the CLARIN infrastructure. In K. De Smedt (Ed.), *Selected Papers from the CLARIN Annual Conference 2015*, 123 (pp. 13–24). CLARIN Annual Conference 2015, 14–16 October 2015, Wrocław, Poland. Linköping University Electronic Press. <https://www.ep.liu.se/ecp/article.asp?issue=123&article=002>
- Kelli, A., Vider, K., Pisuke, H., Lindén, K. (2016). Digitaalsete keeleressursside loomist ja kasutamist määrav õiguslik raamistik Eestis ja selle ühildumine CLARIN-i infrastruktuuriga [Regulatory framework determining the development and utilization of digital language resources and technologies in Estonia and its compability with CLARIN infrastructure]. *Eesti Rakenduslingvistika Ühingu aastaraamat*, 12, 81–98.
- Kõve, V. (2008). Draft common frame of reference and Estonian Law of Obligations Act: Similarities and differences in the system of contractual liability. *Juridica International*, 14, 199–208.
- Law of Obligations Act (2002). Entered into force on 1 July 2002; RT I 2001, 81, 487. <https://www.riigiteataja.ee/en/eli/524012017002/consolide>
- Law of Property Act (1993). Entered into force on 1 December 1993; RT I 1993, 39, 590. <https://www.riigiteataja.ee/en/eli/526012017002/consolide>
- Lehmann, M. (2016). A question of coherence: The proposals on EU contract rules on digital content and online sales. *Maastricht Journal of European and Comparative Law*, 23(5), 752–774.
- Loos, M. B. M., Helberger, N., Guibault, L., Mak, C., Pessers, L., Cseres, K. J., van der Sloot, B., & Tigner, R. (2011). *Analysis of the applicable legal frameworks and suggestions for the contours of a model system of consumer protection in relation to digital content contracts. – Final report: Comparative analysis, law & economics analysis, assessment and development of recommendations for possible future rules on digital content contracts*. University of Amsterdam, Centre for the Study of European Contract Law. <https://hdl.handle.net/11245/1.345662>
- Maňko, R. (2016). *Contracts for supply of digital content: A legal analysis of the proposed new directive*. European Parliamentary Research Service. http://www.europarl.europa.eu/RegData/etudes/IDAN/2016/582048/EPRS_IDA%282016%29582048_EN.pdf
- Maňko, R., & Monteleone, S. (2017). *Contracts for the supply of digital content and personal data protection* [Briefing]. European Parliamentary Research Service. [http://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_BRI\(2017\)603929](http://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_BRI(2017)603929)

- Ministry of Education and Research. (n.d.). *Estonian Language Technology 2018–2027*. <https://www.keeletehnoloogia.ee/en/the-language-technology-research-and-development-program-2018-2027> https://www.keeletehnoloogia.ee/en/projects-2011-2017?set_language=en
- Narciso, M. (2017). Consumer expectations in digital content contracts – An empirical study. *Tilburg Private Law Working Paper Series No. 01/2017*. <https://ssrn.com/abstract=2954491>.
- National Programme for Estonian Language Technology 2011–2017. (n.d.). https://www.keeletehnoloogia.ee/en/projects-2011-2017?set_language=en
- Ohta, T. (2015, January 30). Exhaustion of rights (first sale doctrine): what are the broader implications of the CJEU's ruling in Art & Allposters? *The IPKat*. <https://ipkitten.blogspot.com/2015/01/exhaustion-of-rights-first-sale.html>
- Oprysk, L. (2020). *Reconciling the material and immaterial dissemination rights in the light of the developments under the EU copyright acquis* [Doctoral dissertation]. University of Tartu.
- Ortiz, R. I., & Viscasillas, P. P. (2012). The scope of the common European sales law: B2B, goods, digital content and services. *Journal of International Trade Law & Policy*, 11(3), 241–258.
- Paal, K. (2007). Lease contract, §§ 271–338. In P. Varul, I. Kull, V. Kõve, & M. Käerdi (Eds.), *Võlaõigusseadus II. Kommenteeritud väljaanne [Comments to the Law of Obligations Act, II]*. Juura.
- Petrusson, U. (2011). Patents and open access in the knowledge economy. In A. Kur & V. Mizaras (Eds.), *The structure of intellectual property law. Can one size fit all?* Edward Elgar Publishing.
- Reed, C. (2010). Online and offline equivalence: Aspiration and achievement. *International Journal of Law and Information Technology*, 18(3), 248–273.
- Saare, K., Sein, K., & Simovart, M. A. (2008). The buyer's free choice between termination and avoidance of a sales contract. *Juridica International*, 15, 43–53.
- Sein, K. (2017). What rules should apply to smart consumer goods? Goods with embedded digital content in the borderland between the digital content directive and “normal” contract law. *Journal of Intellectual Property, Information Technology and Electronic Commerce Law*, 8(2), 96–110.
- Smits, J. (2016). The new EU proposal for harmonised rules for the online sales of tangible goods (COM (2015) 635): Conformity, lack of conformity and remedies. *Maastricht European Private Law Institute Working Paper 2016/01*.
- Spindler, G. S. (2016). Contracts for the supply of digital content – Scope of application and basic approach – Proposal of the Commission for a directive on contracts for the supply of digital content. *European Review of Contract Law*, 12(3), 183–217.
- Värv, A., & Karu, P. (2009). The seller's liability in the event of lack of conformity of goods. *Juridica International*, 16, 85–93.
- Yin, R. K. (2009). *Case study research: Design and methods* (4th ed.). SAGE.
- Zoll, F. (2016). The remedies in the proposals of the online sales directive and the directive on the supply of digital content. *European Common Market Law Review*, 6, 250–254.

Copyright © 2020 by author(s) and Mykolas Romeris University

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access